Advances in Water Resources 109 (2017) 196–210

Contents lists available at ScienceDirect

Advances in Water Resources

journal homepage: www.elsevier.com/locate/advwatres

Balancing exploration, uncertainty and computational demands in many objective reservoir optimization

Jazmin Zatarain Salazar[a], Patrick M. Reed[*,a], Julianne D. Quinn[a], Matteo Giuliani[b], Andrea Castelletti[b,c]

[a] School of Civil & Environmental Engineering, Cornell University, Ithaca, NY, USA
[b] Department of Electronics, Information, and Bioengineering, Politecnico di Milano, Milano, Italy
[c] Institute of Environmental Engineering, ETH Zurich, Zurich, Switzerland

ARTICLE INFO

Keywords:
Multi-purpose reservoir control
Direct policy search
Multi-objective evolutionary optimization
Parallel strategies
Uncertainty

ABSTRACT

Reservoir operations are central to our ability to manage river basin systems serving conflicting multi-sectoral demands under increasingly uncertain futures. These challenges motivate the need for new solution strategies capable of effectively and efficiently discovering the multi-sectoral tradeoffs that are inherent to alternative reservoir operation policies. Evolutionary many-objective direct policy search (EMODPS) is gaining importance in this context due to its capability of addressing multiple objectives and its flexibility in incorporating multiple sources of uncertainties. This simulation-optimization framework has high potential for addressing the complexities of water resources management, and it can benefit from current advances in parallel computing and meta-heuristics. This study contributes a diagnostic assessment of state-of-the-art parallel strategies for the auto-adaptive Borg Multi Objective Evolutionary Algorithm (MOEA) to support EMODPS. Our analysis focuses on the Lower Susquehanna River Basin (LSRB) system where multiple sectoral demands from hydropower production, urban water supply, recreation and environmental flows need to be balanced. Using EMODPS with different parallel configurations of the Borg MOEA, we optimize operating policies over different size ensembles of synthetic streamflows and evaporation rates. As we increase the ensemble size, we increase the statistical fidelity of our objective function evaluations at the cost of higher computational demands. This study demonstrates how to overcome the mathematical and computational barriers associated with capturing uncertainties in stochastic multiobjective reservoir control optimization, where parallel algorithmic search serves to reduce the wall-clock time in discovering high quality representations of key operational tradeoffs. Our results show that emerging self-adaptive parallelization schemes exploiting cooperative search populations are crucial. Such strategies provide a promising new set of tools for effectively balancing exploration, uncertainty, and computational demands when using EMODPS.

1. Introduction

Managing river basin systems represents a major global challenge given increasingly uncertain tradeoffs across sectoral uses due to climate change, growing population pressures, and the continued debate over how to sustain environmental services. The computational demands and mathematical difficulty of balancing the inherent multi-sectoral tradeoffs associated with reservoir operations remain a critical research focus recognized by the U.S. Global Research Program workshop (Moss et al., 2016). Emerging computational platforms hold promise for innovating reservoir management by reducing the degree of simplifications and approximations that have traditionally made it difficult to exploit high fidelity simulation models (Giuliani et al., 2015b). As reviewed by Alba et al. (2013), parallel metaheuristics are opening new avenues for simulation-optimization frameworks across a range of application areas. Similarly, recent reviews in the water resources literature (Maier et al., 2014; Nicklow et al., 2009) highlight that advances in parallel computing platforms and meta-heuristics have gained popularity across a wide range of applications (e.g., water distribution systems, groundwater management, surface water management, etc.).

At present, multi-objective evolutionary algorithms (MOEAs) represent one of the fastest growing areas in the water resources planning and management literature. The popularity of MOEAs lies in their ability to provide an explicit understanding of the systems' tradeoffs (Reed et al., 2013), which has been a longstanding focus in water

* Corresponding author.
E-mail addresses: jz583@cornell.edu (J. Zatarain Salazar), pmr82@cornell.edu, patrick.reed@cornell.edu (P.M. Reed).

http://dx.doi.org/10.1016/j.advwatres.2017.09.014
Received 25 May 2017; Received in revised form 11 September 2017; Accepted 14 September 2017
Available online 18 September 2017
0309-1708/

management (Cohon and Marks, 1975; Haimes and Hall, 1977). MOEAs exploit population-based search to discover the full set of Pareto approximate solutions in a single run (see review in Coello et al. (2007)). The Pareto-approximate solutions are those for which improvement in one objective can only be achieved by sacrificing performance in one or more other objectives. MOEAs support simulation-optimization applications with challenging mathematical properties such as non-convexity, nonlinearity, stochasticity, mixtures of continuous and discrete decision variables, multimodality, and stochastic evaluations (Coello et al., 2007; Reed et al., 2013). It is increasingly more common in MOEA applications to exploit Monte Carlo (MC) simulations to discover robust operating policies (Hamarat et al., 2014; Mortazavi-Naeini et al., 2015; Müller and Schütze, 2016; Tsoukalas et al., 2016; Zhang et al., 2017).

Despite their advantages, the simulation-based search required by MOEAs can still pose high computational costs. Given their stochastic nature, MOEAs need to be run under multiple random seed trials to account for performance differences that may emerge for different randomly generated initial populations or across randomized sequences of their search operators. These trials pose an additional computational demand that can limit the size or scope of applications. Additionally, the growing prevalence of using MC simulations to evaluate uncertain objectives also increases computational demands and adds mathematical difficulty to the search problems (Reed et al., 2013; Singh and Minsker, 2008; Ward, 2015). Several studies have highlighted an inherent tension between the level of approximation in MC-based evaluations of objectives (i.e., small versus large sampling rates) and the degree to which an MOEA is able to explore the decision space (i.e., the number of function evaluations, NFE) (Gopalakrishnan et al., 2003; Kasprzyk et al., 2013; Nicklow et al., 2009; Singh and Minsker, 2008). Less approximate MC evaluations of objective functions are computationally demanding, thus limiting the scope of search in a given period of wall-clock time. However, faster but more approximate MC evaluations can mislead MOEAs into mistakenly classifying some solutions as superior to others in all objectives. Such ranking errors can prematurely eliminate important solutions simply as an artifact of noisy evaluations (see detailed discussions in Deb and Gupta (2006) and Beyer and Sendhoff (2007)).

Current innovations in parallel MOEA strategies can help overcome the computational bottleneck posed by expensive MC-based function evaluations while also improving the efficiency, effectiveness and reliability of MOEAs when dealing with noisy optimization (Alba et al., 2013; Maier et al., 2014; Reed and Hadka, 2014; Tang et al., 2007). Efficiency refers to minimizing the computational time to attain solutions. Effectiveness requires high quality Pareto approximation sets. Algorithmic reliability seeks to maintain high levels of efficiency and effectiveness across all random trials. Recent studies (Reed and Hadka (2014) and Giuliani et al., 2017) have made significant efforts to assess the scalability of massively parallel runs that demonstrate the potential of MOEAs for large-scale water management applications. We expand their previous efforts by comprehensively diagnosing the ability of two parallel variants of the self-adaptive Borg MOEA to support many objective reservoir optimization under hydro-climatic uncertainty. The master-worker and multi-master worker variants of the Borg MOEA are evaluated in this study for their ability to find Pareto approximate control policies for the Lower Susquehanna River Basin (LSRB).

The LSRB system needs to meet several competing water demands for hydropower production, recreation, drinking water distribution for Baltimore, MD and Chester, PA, cooling water for the Peach Bottom Atomic Power Plant, and federally mandated environmental flows for fish passage. We build our work on earlier efforts by Giuliani et al. (2014) to effectively capture the LSRB's tradeoffs using Evolutionary Multi-objective Direct Policy Search (EMODPS) (Giuliani et al., 2015a). EMODPS is a simulation-based approach to discover Pareto approximate control policies for multi-purpose reservoir operations. With EMODPS, the control policy is first parameterized using global approximators (i.e., radial basis functions or artificial neural networks) to provide more flexibility to the shape of the control policy. The parameters are then optimized with respect to the objective functions using parallel Borg MOEA variants.

The LSRB problem represents a different class of problems related to many-objective stochastic control than the risk-based water supply portfolio planning problem for which Hadka and Reed (2015) performed a diagnostic assessment of alternative Borg MOEA parallelizations. The LSRB captures a suite of difficult mathematical properties including heterogeneity in the mathematical formulation of objectives (expectation over time and minmax over ensembles), a large number of objectives, and diverse decision time-scales across the sectoral interests. Furthermore, we benchmark the performance of alternative parallelization schemes of the Borg MOEA with tens to hundreds of compute cores, a size more representative of commonly available high-performance computing platforms relative to our prior work exploring 10^3 to 10^5 core systems (Hadka and Reed, 2015; Reed and Hadka, 2014). In summary, this study requires that the advanced parallel instances of the Borg MOEA solve a highly demanding stochastic and many-objective reservoir control problem with 2 or 3 orders of magnitude as many compute cores as in our prior published benchmarking applications (i.e., 32–256 cores).

In addition to benchmarking the performance of different parallel versions of the Borg MOEA on a smaller cluster, in this study we also analyze the tradeoff between better representing uncertainty in the stochastic optimization and simplifying this representation to shorten the function evaluation time and allow for greater search by optimizing reservoir operations to different size ensembles of stochastic hydrologic inputs. To our knowledge, this is the first study to bridge analyses of the tradeoff between communication and depth of searc with analyses of the tradeoff between model fidelity and depth of search. Importantly, we find that cooperative parallel strategies of the Borg MOEA enable high quality representation of tradeoffs as we transition to more challenging(higher fidelity)formulations of the LSRB with better representations of hydro-climatic uncertainties. We show for the first time that even on a relatively modest scale cluster system (i.e., hundreds of computing cores), the cooperative Borg MOEA parallelization scheme is able to better cope with the high dimensional LSRB stochastic control problem and its expensive function evaluations while improving the accuracy of its estimated Pareto set within a limited search time, driving the front towards a better approximation with improved convergence speed and reliability.

This paper is organized as follows: in Section 2 we provide a description of the LSRB system, the model formulation and the synthetic hydrology generation. In Section 3, we detail the EMODPS implementation for the LSRB, describe the parallel Borg MOEA strategies and define the performance metric used to measure the quality of the Pareto-approximations. In Section 4 we present the computational experiment. In Section 5 we present the results of effectiveness, reliability, efficiency and algorithmic speedup of each configuration tested compared against individual MC sampling schemes, and we visualize the tradeoffs obtained across all configurations and samples tested.

2. The Lower Susquehanna River Basin system

The LSRB system serves several competing demands that are common to reservoir systems worldwide: (1) revenues from hydropower production, (2) cooling water for major energy systems, (3) urban water supply, (4) recreational uses, and (5) regulated environmental flow requirements. These multi-sectoral demands must be balanced while appropriately addressing hydro-climatic variability and evolving societal values (Sheer and Dehoff, 2009). The Conowingo Reservoir, located in the LSRB, as seen in Fig. 1, is an interstate water body with regional water supply pump stations for Chester, PA and Baltimore, MD. The Conowingo dam was completed in 1968 for

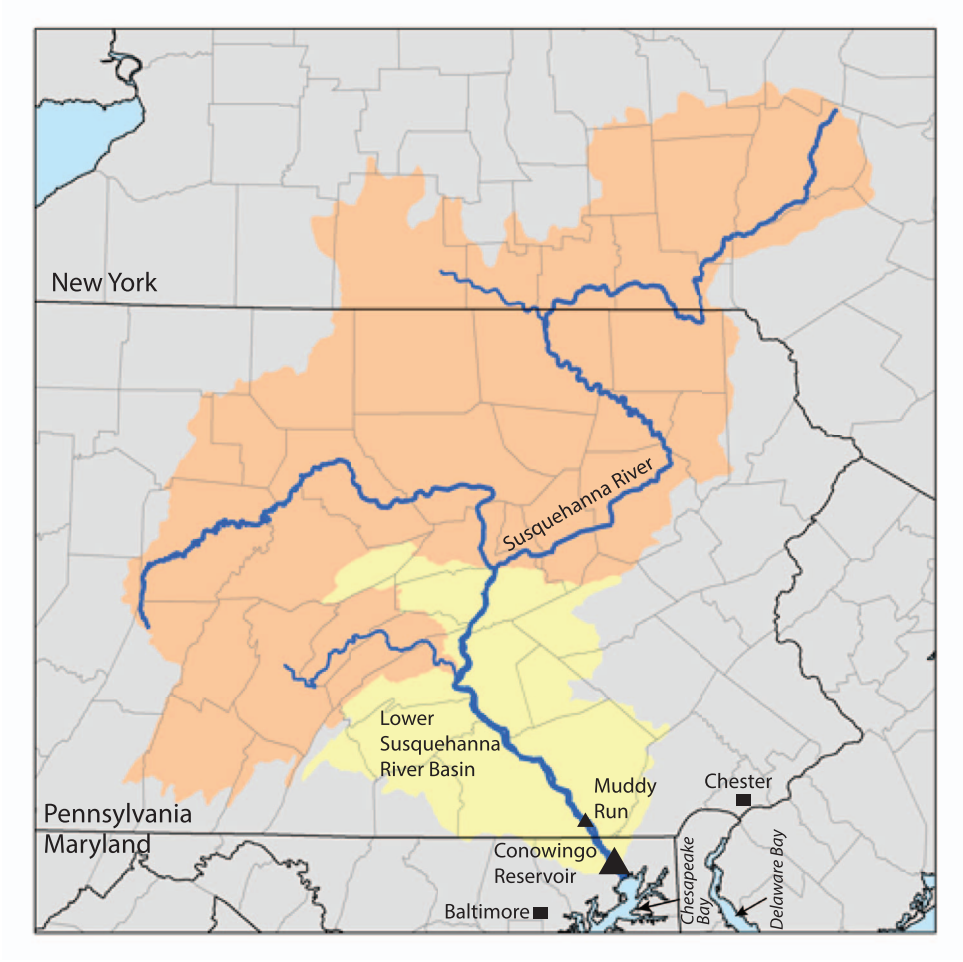

Fig. 1. Map of the Susquehanna River Basin, highlighted in yellow is the Lower Susquehanna River sub-basin. The Conowingo hydropower plant (large triangle) and the Muddy Run pumped-hydro (smaller triangle) facility are located between Pennsylvania and Maryland (Map adapted from Musser 2007). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

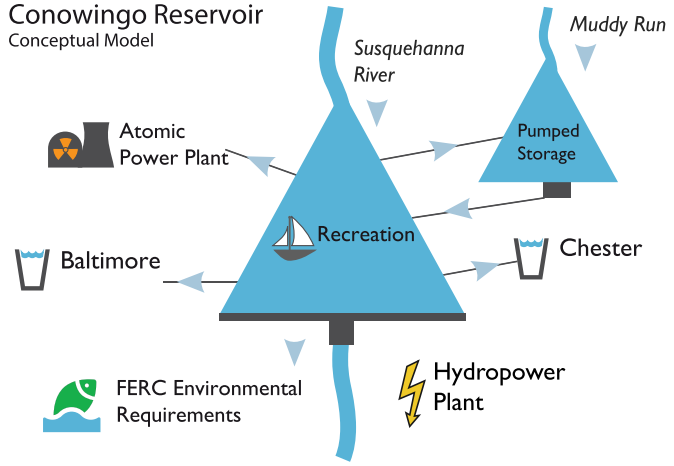

Fig. 2. Schematic representation of the Conowingo Reservoir mass balance model and the six objectives.

hydropower production with a total capacity of 548 MW and is the largest privately owned dam in the U.S. Reservoir operations at the Conowingo dam regulate 50% of the freshwater flows going into the Chesapeake Bay. The reservoir has a total capacity of 310,000 acre-ft (0.38 km^3) and plays a key thermal regulation role. The Conowingo reservoir provides cooling water for the Peach Bottom Nuclear Generating Station. Additionally, reservoir operations are used to maintain federally mandated minimum flow requirements set by the Federal Energy Regulatory Commission (FERC) to allow fish passage during the migratory season. Mediating these multi-sectoral demand tradeoffs has resulted in the LSRB being adaptively managed by the Susquehanna River Basin Commission (SRBC) through computer aided negotiations with the system's regional stakeholders (Sheer and Dehoff, 2009).

2.1. Conowingo Reservoir simulation

This work broadens the space of candidate operations and objectives that can be evaluated relative to prior SRBC computer aided modeling efforts (Randall et al., 1997; Sheer and Dehoff, 2009; Swartz, 2006). The simulation model for the LSRB system is based on the prior development efforts of Giuliani et al. (2014). We simulate the Conowingo system using a series of stochastically generated stream inflows and evaporation rates that compensate for the limited duration of historical observations, better capturing the variability of extreme flood and drought conditions (see Section 2.2 below). The mass balance model is based on the interaction between the Conowingo Reservoir and the Muddy Run Pumped Storage Hydroelectric Facility, which takes advantage of intra-daily cycles in energy prices. The system uses excess power in the grid to pump water into Muddy Run during off-peak hours, and then returns the water by gravity to Conowingo during peak hours (Fig. 2).

The dynamics between both reservoirs are described by Eqs. (1) and (2).

$$s_{t+1}^{CO} = s_t^{CO} + q_{t+1}^{CO} + q_{t+1}^{CO,L} - r_{t+1}^{CO} - E_{t+1}^{CO} - q_{t+1}^{p} + r_{t+1}^{MR} \quad (1)$$

$$s_{t+1}^{MR} = s_t^{MR} + q_{t+1}^{MR} - r_{t+1}^{MR} - E_{t+1}^{MR} + q_{t+1}^{p} \quad (2)$$

where s_t^{CO} and s_t^{MR} are the volumes of water stored at Conowingo (CO)

and Muddy Run (MR), respectively, and $q_{t+1}^{CO,L}$ are the mainstem and lateral inflows measured at the Marietta gauging station, q_{t+1}^{P} is the water pumped from Conowingo to Muddy Run, and q_{t+1}^{MR} is the inflow to Muddy Run. E_{t+1}^{CO} and E_{t+1}^{MR} are the evaporation losses measured at each reservoir, equal to the average evaporation rates during a 4-h simulation step at time $(t + 1)$, e_{t+1}^{CO} and e_{t+1}^{MR}, multiplied by the surface area of the Conowingo and Muddy Run reservoirs at time $(t+1)$, respectively. The released volume is given by $r_{t+1}^{i} = f(s_t^i, u_t^i, q_{t+1}^i, E_{t+1}^i)$ which depends upon the storage s_t^i, the policy-prescribed release decision u_t^i, the inflow q_{t+1}^i, and the evaporation loss E_{t+1}^i, where the superscript i refers to either Conowingo or Muddy Run (Soncini-Sessa et al., 2007). The time subscript refers to the time instant at which it assumes a deterministic value. The reservoir storage is measured at time t, whereas the inflows are measured in the time interval $[t, t + 1)$. The 4-hour decision time step was used in order to have a time step small enough to account for fluctuations in energy prices, and large enough to avoid the impact of turbine dynamics (e.g. cavitation and ramp up).

2.2. Stochastic hydrology

In this study, statistical synthetic hydrology is used to better capture the LSRB system's flood and drought extremes, given that these conditions are rarely observed in the historical streamflow records. Consequently, solely considering historical hydrology in the formulation and evaluation of the Conowingo Reservoir's operating policies would systematically underestimate the impacts of hydrologic variability and extremes (Loucks et al., 2005). The core mathematical benchmarking challenge in this study is to find control policies and their tradeoffs that emerge across the uncertain and highly variable streamflow conditions common to the LSRB. Synthetic hydrologic ensembles provide a means of including MC simulation of hydro-climatic uncertainties while maintaining key temporal and spatial statistical traits.

The LSRB synthetic hydrologic ensembles were generated on a monthly time step using the method of Kirsch et al. (2012) and then disaggregated to daily values using the method of Nowak et al. (2010). The method of Kirsch et al. (2012) preserves autocorrelation at the monthly time step for a particular hydrologic variable by multiplying a matrix **C** of randomly sampled historical observations by an upper triangular matrix obtained from Cholesky decomposition of that variable's historical autocorrelation. Cross correlation between the hydrologic variables (i.e., flows at the Marietta gauging station, lateral inflows between Marietta and Conowingo Dam, inflows to Muddy Run, and evaporation rates at Conowingo and Muddy Run dams) is preserved by using historical values from the same historical month to populate **C** for each variable. Using the method of Nowak et al. (2010), the hydrologic variables are disaggregated to daily values by randomly selecting one of k nearest neighbors from the historical record in terms of total monthly flows/evaporation rates across sites and proportionally scaling that month's daily values at each site to match the synthetic monthly total for each variable.

This method has been shown to successfully replicate the seasonal correlation structure while producing extreme flow events outside of those observed in the historic record (Herman et al., 2016; Kirsch et al., 2012). A more detailed mathematical description of the synthetic generation and statistical validation of the synthetic hydrologic variables is provided in the supplement.

Historical LSRB streamflows from 1932 to 2001 were used to build the synthetic generator to simulate time-series of the main and lateral streamflow to Conowingo (q_{t+1}^{CO},$q_{t+1}^{CO,L}$), as well as the Muddy Run streamflow (q_{t+1}^{MR}). Similarly, we simulated evaporation rates at Conowingo (e_{t+1}^{CO}) and Muddy Run (e_{t+1}^{MR}) where $e_{t+1}^{CO} = e_{t+1}^{MR}$. A total of 10,000 annual traces were generated for each of the above-mentioned variables. Fig. 3 panel (a) shows the annual historical flow duration curves (in blue) versus the stochastically generated flow duration curves (in gray) for that provide a broader sample of extreme drought and flood conditions in comparison to the historical flows. Relative to synthetic hydrologic scenarios used in Giuliani et al. (2014), this study added a joint resampling scheme to better capture the spatial correlation between Muddy Run's lateral flows and the Conowingo reservoir's inflows.

As noted in our Introduction, a core challenge and contribution in this study is addressing the complex balance between the computational demands posed by MC-based solution evaluations versus the impacts of highly approximate but fast evaluations. For the LSRB model, we explore these issues by including three levels of approximation for the MC evaluations: 50, 500 and 1000 ensemble members. The synthetic hydrologic series used for each MC sampling level were selected from the 10,000 stochastic samples based on each trace's annual flow at Marietta, shown in Fig. 3 panel (a). The ensemble members with the wettest and driest years at the Conowingo reservoir were included in each ensemble (i.e. the years with the highest and lowest annual streamflow) with the remaining years selected from a one-dimensional Latin hypercube sample of the inverse empirical cumulative distribution of annual Marietta flows, . For each selected synthetic Marietta flow series, the corresponding evaporation rate and inflow series associated with that trace were also used. The flow duration curves of the three approximate MC sampling levels are shown in panel (b) of Fig. 3.

3. Methods

The Conowingo reservoir represents an ideal test case given the array of conflicting multi-sector demands and significant hydro-climatic variability that characterize the LSRB system. The Conowingo reservoir's operations are explored in this study using the Evolutionary Multi-Objective Direct Policy Search (EMODPS) framework (Giuliani et al., 2015a). The EMODPS framework provides a generic approach for discovering Pareto approximate reservoir control policies that seek to balance conflicts across competing demands. As discussed in the Introduction, EMODPS is a simulation-based approach for discovering candidate control policies. The operating policies are first parameterized using some form of mathematical functions (e.g., piecewise linear functions, radial basis functions, neural networks, etc.) and then the parameters of those functions are optimized with respect to the reservoir's operating objectives using MOEAs. The term direct policy search (DPS) was first introduced by Rosenstein and Barto (2001) in the control literature. Independently, the same concepts have been termed parameterization-simulation-optimization in the water resources literature (Koutsoyiannis and Economou, 2003), with early example applications dating back to the 1980s (Guariso et al., 1986; Oliveira and Loucks, 1997).

One of the most active areas of DPS research lies in extending the framework to multiple objectives. MOEAs are particularly attractive for DPS applications given the potential nonlinearity of candidate control policies, the complexities of capturing uncertainties, and multi-sector tradeoffs (Giuliani et al., 2017). Our previous diagnostic study (Zatarain Salazar et al., 2016) assessed the capability of state-of-the-art MOEAs to support DPS for many-objective reservoir control where the Borg MOEA demonstrated potential for benefiting from parallelization strategies by having continued progress throughout the search, and performing consistently well after a given NFE count. In this study, we contribute a diagnostic study to benchmark different parallelization strategies of the Borg MOEA to discover tradeoffs and support decision making for computationally intensive many-objective reservoir control problems under hydro-climatic uncertainties. Section 3.1 presents the control policy formulation. Section 3.2 overviews the objective function formulations. Section 3.3 presents the Borg MOEA parallel variants. Section 3.4 describes the metrics used to test algorithmic efficiency and solution quality.

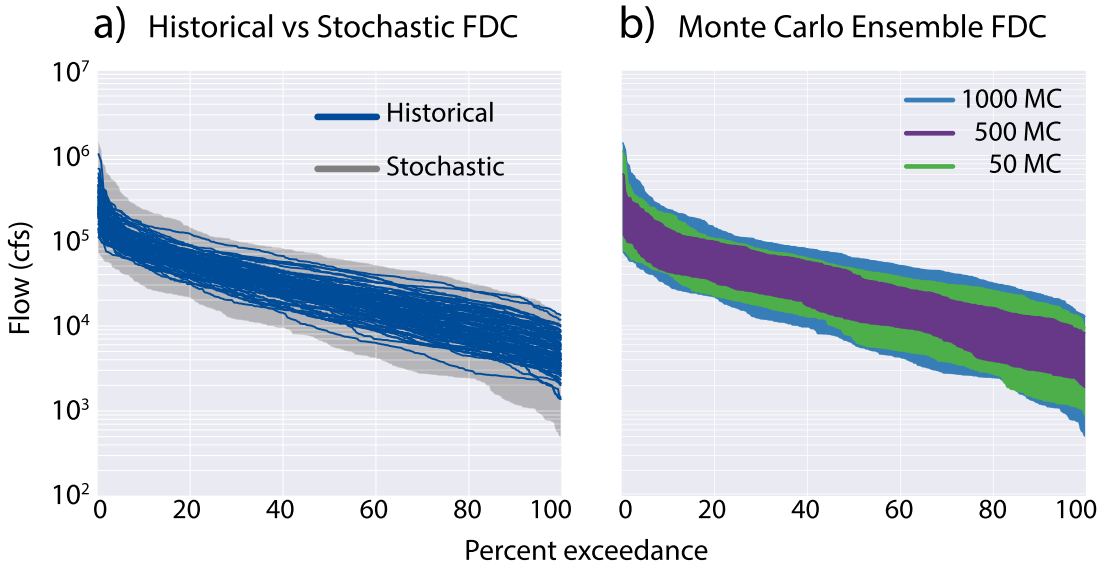

Fig. 3. Panel (a) shows the annual flow duration curves (FDC) of the flow at the Marietta gaging station. The observed record (1930–2012) is in blue and the stochastic ensemble is in gray. Panel (b) shows annual flow duration curves sampled from the stochastic ensemble, preserving the wettest and driest years across the ensemble. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3.1. Control policy formulation

In this study, we parameterize the control policies using radial basis functions to map the reservoir level and time index into release decisions. The release from Conowingo (r_{t+1}^{CO}) is actually a vector of 4 releases corresponding to the water supply for Chester (r_{t+1}^{C}), Baltimore (r_{t+1}^{B}), the Atomic Power Plant (r_{t+1}^{APP}), and the downstream release for environmental flows and hydropower production (r_{t+1}^{EHP}). The k^{th} policy-prescribed release decision (u_t^k) in the time interval $[t, t+1)$ is as follows:

$$u_t^k = u_\theta^k(z_t) = \sum_{i=1}^{n} w_i^k \varphi(z_t) \tag{3}$$

where n, is the number of RBFs, w_i^k is the weight of the ith basis function for the kth release, and z_t is the input vector, where $\mathbf{z_t} = |z_{t,1}, \ldots, z_{t,m}|$. Each Gaussian RBF (φ) is defined by Eq. (4):

$$\varphi_i(z_t) = exp\left[-\sum_{j=1}^{m} \frac{(z_{t,j} - c_{j,i})^2}{b_{j,i}^2}\right] \tag{4}$$

where m is the number of input variables, and c_{ij} and b_i are the centers and radii, respectively, of the ith RBF for the jth input variable. Here we used $n = 4$ RBFs, with the parameter vector defined as $\theta = |c_{i,j}, b_{i,j}, w_i^k|$, with $i = 1, \ldots, n, j = 1, 2$ corresponding to the time index and the reservoir level, and $k = 1, \ldots, 4$ for the 4 release decisions, yielding a total of 32 policy parameters to be optimized by the MOEA. The inputs in z_t are uniformized on [0, 1], while $c_{i,j} \in [-1, 1]$, $b_{i,\ j} \in [0, 1]$ and $w_i^k \in [0, 1]$ with $\sum_{i=1}^{n} w_i^k = 1$.

3.2. Formulation of objectives

The performance of the policy parameters (θ) is evaluated in the objective space by solving the following multi-objective problem:

$$\theta^* = arg \min_{\theta} J(\theta) \tag{5}$$

The policy parameters θ^* are the Pareto approximate decision variables found by the MOEA. The parameters are obtained by simulating the system over a time horizon of one year under the policy $p = \{u_{\theta^*} = |u_{\theta^*}^C, u_{\theta^*}^B, u_{\theta^*}^{APP}, u_{\theta^*}^{EHP}; t = 0, \ldots, H-1\}$ where H is the simulation horizon. The objective functions J are the reservoirs operating objectives, defined by the following objective formulations:

Hydropower revenue (to be maximized) is defined as the economic revenue obtained from hydropower production at the Conowingo hydropower plant (hyd) provided by Eq. (6).

$$\Phi_{hyd} = \sum_{t=1}^{H} (HP_t \cdot \rho_t) \tag{6}$$

Revenue is a function of the hourly energy production (HP_t) in MWh and the hourly energy price (ρ_t) in USD/ MWh, defined by a seven hour moving average of the energy price trajectory in the Pennsylvania, New Jersey–Maryland (PJM) energy market. The hourly energy production in MWh is defined as:

$$HP_t = \eta g \gamma_w \overline{h}_t q_t^{Turb} \cdot 10^{-6} \tag{7}$$

where η is turbine efficiency, g is the gravitational acceleration (9.81 m/s^2), γ_w is the density of water (1000 kg/m^3), h_t is the net hydraulic head in meters, $q^{Turb} = min(r_t^{EHP}, q^{\max})$ is the turbined flow in m^3/s, equal to the minimum of the downstream release for environmental flows and hydropower production (r_t^{EHP}), and the maximum capacity of the turbines ($q^{\max}$).

Water supply reliability to Baltimore, Chester and the Atomic Power Plant (to be maximized) is measured by the daily average volumetric reliability (VR) in Eq. (8) (Hashimoto et al., 1982).

$$\Phi_{VR,s} = \frac{1}{H} \sum_{t=1}^{H} \frac{Y_t^s}{D_t^s} \tag{8}$$

where Y_t^s is the daily delivery in m^3, D_t^s is the daily demand in m^3, and the subscript s corresponds to the water supply to either Baltimore, Chester or to the Atomic Power Plant (Fig. 2).

Recreation (to be maximized) is defined as the storage reliability (SR) on weekends of the touristic season described by the relationship between the number of weekend days in the touristic season below the target level (n_F) and the total number of weekends in the touristic season (N_{we}) during which the target level is 106.5ft (32.5 m) to guarantee boating.

$$\Phi_{SR} = 1 - \frac{n_F}{2N_{we}} \tag{9}$$

Environmental shortage (to be minimized) is defined as the daily average shortage index (SI) relative to the FERC flow requirements described in Eq. (10). A quadratic function is used in order to penalize larger deficits while allowing small and more frequent shortages

(Hashimoto et al., 1982).

$$\Phi_{SI} = \frac{1}{H}\sum_{t=1}^{H}\left(\frac{max(Z_t - r_t^{EHP}, 0)}{Z_t}\right)^2 \quad (10)$$

Z_t and r_t^{EHP} are the FERC flow requirement and the daily environmental release, respectively, both in m^3.

The objectives where evaluated over synthetic realizations of inflow and evaporation rates. A minimax approach formulation was adopted where the objectives are minimized in the worst-case realization. The minimax operator was independently applied for each objective (J^k), providing a sense of their lower bound performance shown in Eq. (11) below:

$$J^k = \Psi_{i\in(1,\ldots,N)}[\Phi_{t\in(1,\ldots,H)}[g^k(t,i)]] \quad (11)$$

where g^k is the value of the k^{th} objective during simulation period t in ensemble member i, with $k = hyd, VR, SR, SI$. Φ is the aggregation of the objectives $[g^k(t,i)]$ over the 1-year simulation horizon (H) as shown by Eqs. (6), (8)–(10), and Ψ is the filtering operator across N ensembles.

The minimax nature of the objective functions in Eq. (11), make performance evaluations strongly sensitive to MC sampling levels. The three approximation levels for MC sampling explored in this study influence the mathematical difficulty of the optimization as well as the overall computational demands; as the MC sample size increases it becomes harder to find solutions that perform well across the larger ensemble and the computational time per function evaluation increases.

3.3. Parallelization schemes of the Borg MOEA

The serial version of the Borg MOEA (Hadka and Reed, 2012, 2013) incorporates key features that have proven to contribute to successful search: (1) the use of epsilon dominance-archiving (Laumanns et al., 2002); (2) auto-adaptive population sizing and selection; (3) dynamic adaptation and cooperative use of multiple candidate mating and mutation operators; and (4) auto-adaptive triggering of a re-start mechanism that combines the search archive and uniform mutation to diversify search around the best known solutions when progress stalls. This study compares two parallelization strategies for the Borg MOEA, the simple master-worker scheme and a more recently contributed multi-master scheme (Hadka and Reed, 2013, 2015). Each of the parallel versions of the Borg MOEA is described in more detail below.

3.3.1. Master-worker Borg MOEA

The master-worker Borg MOEA illustrated in Fig. 4, is the simplest parallelization strategy for the algorithm. In the master-worker Borg MOEA only the function evaluations are parallelized; the internal search dynamics of the original serial algorithm are not modified. Consequently, master-worker strategies simply provide the ability to compute more function evaluations in a fixed wall-clock period as there is no interaction between workers (Tang et al., 2007). In a system with C cores, one of the cores is the master and the remaining C-1 cores are labeled as workers. The master node runs the serial MOEA and dispatches the function evaluations to one of the available worker nodes. The master sends the vector of decision variables to an available worker node, the worker node evaluates the problem with the given decision variables, and when finished, sends the evaluated objective values back to the master node.

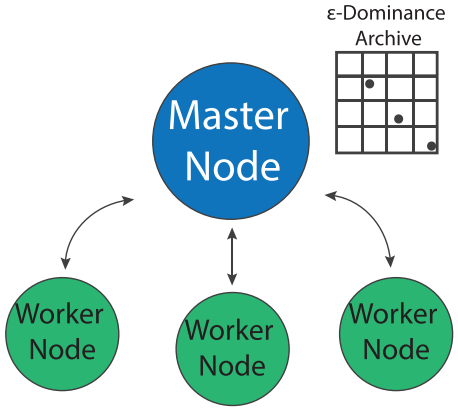

Fig. 4. Master-worker implementation of the Borg MOEA. The master node performs the main serial loop and dispatches the function evaluations to the available worker nodes which then return the evaluated objective functions to the master node, which keeps track of the epsilon-dominance archive. (Adapted from Hadka and Reed (2015)).

While the master-worker parallelization scheme will result in the same search dynamics as the serial algorithm, it should be noted that the time savings associated with parallelization can allow for significantly more search given the Borg MOEA utilizes time-continuation of search, (see Goldberg (2002)). This is especially important for searching water resources applications under uncertainty where MC-based function evaluations can pose significant computational barriers. As mentioned earlier, users must balance the scope of possible search (i.e., total NFE) versus the level of approximation used in MC function evaluations (i.e., the number of samples per evaluation). The master-worker Borg MOEA provides a guarantee of scalably increasing the number of function evaluations that can be performed in a given period of unit wall-clock and of maintaining diverse search. Unfortunately, the parallelization scheme does not significantly change the reliability of the algorithm in confronting mathematically challenging problems that induce search failures in the serial algorithm.

3.3.2. Multi-master Borg MOEA

Beyond simple master-worker schemes, the other most popular implementation of parallel meta-heuristic search has been the multiple search population model, also termed islands or multi-population (see Cantu-Paz (2000)). In this approach, each computing core has a fully functional copy of the search algorithm that interacts with other search populations through cooperative strategies for migrating solutions. The multi-population parallelization scheme has been shown to fundamentally change the meta-heuristics core search dynamics and makes it possible to solve more challenging mathematical problems (see Alba et al. (2013), Tang et al. (2007) and Cantu-Paz (2000); Crowl (1994)). Relative to the master-worker, the two largest drawbacks of the multi-population approach are (1) the reduced granularity of its parallelism limits the scope and predictability of maximizing the NFE that can be considered in a given period of wall-clock time due to increased communication and (2) solution sharing or migration schemes add another level of algorithmic complexity that can make it difficult for users to rapidly and reliably maximize search capabilities.

As illustrated in Fig. 5, Hadka and Reed (2015) developed the multi-master Borg MOEA as a hybrid parallelization scheme that combines the master-worker and multi-population approaches. The multi-master Borg MOEA combines the strengths of both parallelization schemes and generalizes the auto-adaptivity of the algorithm to a highly flexible co-evolutionary framework that can be adapted to a broad range of computing architectures. The multi-master Borg MOEA takes advantage of the cooperative search of distributed populations while maximizing the granularity of its parallelism using worker cores. In its lower bound instantiation, the multi-master Borg MOEA consists of at least two co-evolving and interacting master-worker instances. Each master independently maintains the self-adaptivity of the serial and single-master implementations of Borg but with the ability to communicate with other masters through a controller in order to maximize performance. The controller introduced by the multi-master Borg MOEA (depicted in Fig. 5) has two tasks: (1) maintain a global ϵ-dominance archive, and (2) provide guidance to master nodes when they need help. The global ϵ-dominance archive maintains the Pareto optimal solutions discovered by all masters. The controller uses the global ϵ-dominance archive to track the operators that contribute globally

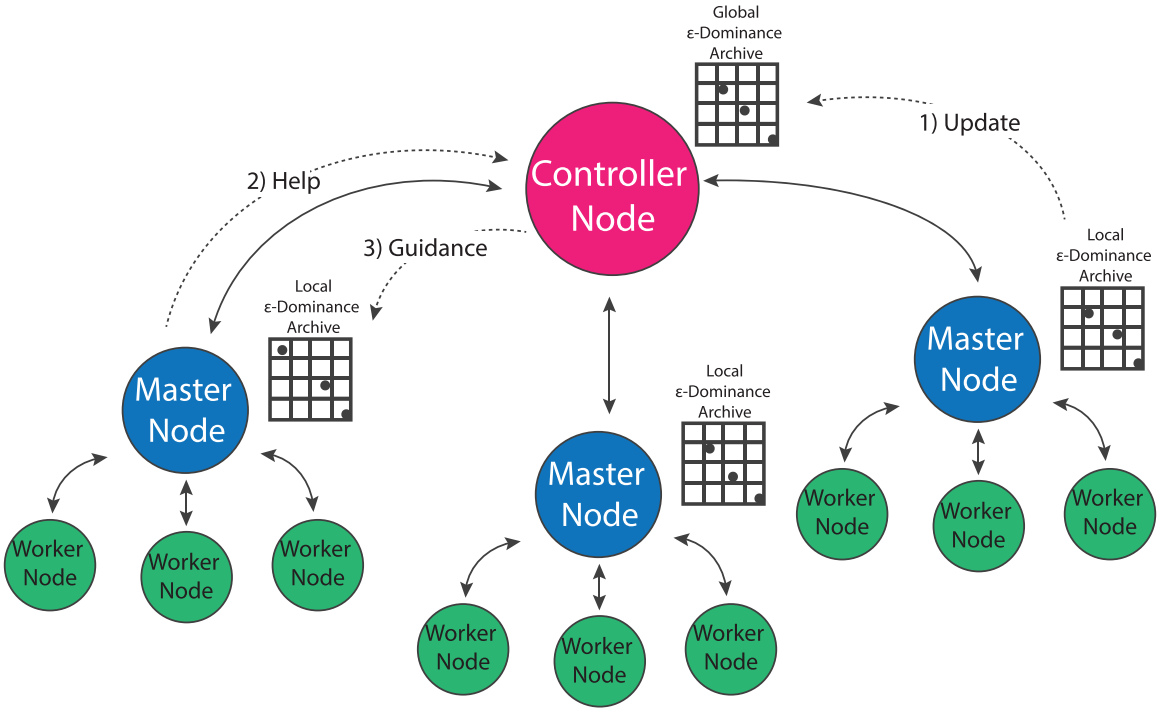

Fig. 5. Diagram illustrating the multi-master Borg MOEA implementation consisting of two or more master-worker instances. The controller node communicates with the masters in three different ways 1) the masters send periodical updates to the controller of their local epsilon dominance archive; 2) when a master node struggles it sends a help request to the controller, which in turn 3) sends guidance, which includes the global epsilon-dominance archive and operator probabilities (Adapted from Hadka and Reed (2015)).

Pareto approximate solutions. Each master node periodically sends an update to the controller. This update contains any new Pareto approximate solutions discovered by the master since its last update and the current probability of using each operator.

Since each master node is running an instance of the master-worker Borg MOEA, it includes all of the mechanisms to detect search stagnation and trigger restarts. If these mechanisms are unsuccessful at escaping the local optima, the master node notifies the controller that it needs help. Once receiving the help request, the controller seeds the master with the global ϵ-dominance archive and global operator probabilities that have been successful in contributing new Pareto approximate solutions so far. Upon receiving this guidance from the controller, the master updates its internal state and triggers a restart.

Additionally, the multi-master implementation also features a different style of initialization from the serial and master-worker Borg MOEA implementations. Instead of uniformly sampling the decision variables of the Borg MOEA at random from within their bounds to generate the initial population, the master-worker and the multi-master worker implementations used in this study utilize an improved statistical initialization by using a Latin Hypercube sample (LHS) to help ensure that the initial population contains a well-dispersed sampling of the search space.

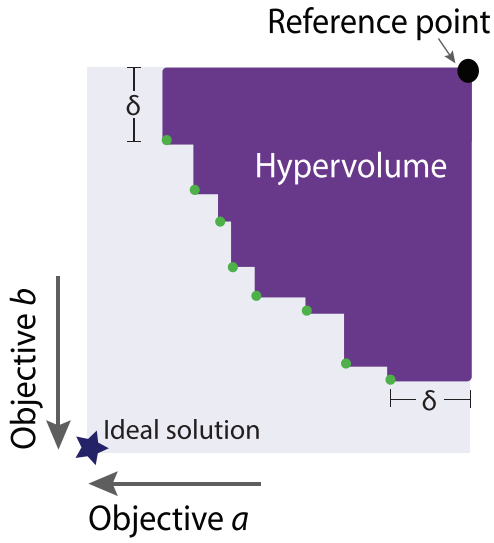

Fig. 6. Schematic of the hypervolume indicator in a 2D projection. The bounds of the reference approximation set are used to calculate the reference point; this calculation typically adds a delta (δ), so that the boundary points contribute positive hypervolume.

3.4. Measuring the quality & speed of search

An effective MOEA must generate approximation sets that are both proximate (i.e., converge to the true Pareto front) and diverse (i.e., capture the geometry and extent of tradeoffs). We use hypervolume (Knowles and Corne, 2002; Zitzler et al., 2003) a metric that captures both diversity and proximity, to distinguish how well alternative parallel Borg MOEA configurations capture the LSRB's tradeoffs. Fig. 6 provides an illustrative example of how hypervolume is computed for a 2-objective problem. A reference point is chosen based on the bounds of the approximation set plus an additional delta. This delta ensures the boundary points contribute positive volume to the overall hypervolume. This metric measures the volume of the objective space dominated by an approximation set. A large hypervolume will correspond to approximation sets that dominate more space, indicating high-quality approximation sets (i.e., proximity and diversity).

In addition to using the hypervolume metric to measure solution quality, we measure parallel performance through speedup. Conventionally speedup is defined as the time required to solve an application serially divided by the time required to solve the same application in parallel. However, in the MOEA search context, speedup must be carefully tied to the quality of the search process itself (i.e., hypervolume). This study uses a hypervolume-based speedup analysis to characterize the relative benefits of the multi-master Borg MOEA versus the simpler master-worker variant of the algorithm. We compute

a relativistic hypervolume speedup using Eq. (12). For each level of hypervolume, the algorithmic speedup (S_A) is computed as the ratio of the NFE required by the master-worker (NFE_{mw}) relative to the NFE required per master in the multi-master scheme (NFE_{mm}). In this experiment, the parallel Borg variants are racing to attain high levels of hypervolume performance (HV) within a fixed wall-clock time (T) using the same number of compute cores (C). Any perceived speedups are then the result of improved algorithmic search.

$$S_A = \left(\frac{NFE_{mw}}{NFE_{mm}}\right)_{T,P} \tag{12}$$

The speedup scalability of the Borg MOEA parallel implementations (or any parallel MOEA) is a function of a complex mixture of factors including the ratio of function evaluation time to communication time, bottlenecks in processing evolutionary operators, and the stochastic loading dynamics of the architecture being used. Readers interested in a more technical discussion and analysis of these factors for the Borg MOEA can reference (Hadka et al., 2013; Reed and Hadka, 2014).

4. Computational experiment

The computational experiment developed for this study evaluates how well each candidate parallel configuration of the Borg MOEA is able to approximate the LSRB's tradeoffs. Fig. 7 provides an illustrative overview of our diagnostic evaluation framework. We ran the experiment on The Cube cluster at the Cornell Center for Advanced Computing, with 32 compute nodes with Dual 8-core E5-2680 CPUs at 2.7 GHz, 128GB of RAM, and exploited Open Message Passing Interface version 1.6.5 in the parallelized Borg MOEA instances.

We used three MC sampling levels to define the Conowingo reservoir's tradeoffs, calculating objective values across 50, 500 and 1000 independent 1-yr ensemble members shown in panel (a). These simulations were run over different configurations of the master-worker and multi-master Borg MOEA implementations, with core counts ranging from 32 to 256, and master count ranging from 1 to 16 (panel b). The single master-count is equivalent to the master-worker configuration, which serves as our speedup baseline shown in Eq. (12). Each parallel Borg MOEA configuration was run for 25 random seed trials to account for variability in their initial populations and operator probabilities, each seed was allocated a 10 h wall-clock time for search. All algorithm configurations were initialized using Latin hypercube samples where each initial Borg MOEA population size was set equal to 100 individuals. The default crossover, mutation and selection parameters of the Borg MOEA used in this study are specified in Table 1.

A key difference between the Borg MOEA parallelization schemes is that the master-worker's evolutionary processes start from a single initial population while in the multi-master scheme the Latin Hypercube samples are distributed across interacting populations. As discussed in Section 3.3, a key algorithmic innovation that enables the Borg MOEA to maintain diverse approximations to the Pareto set is the ϵ-dominance archiving. For the Susquehanna objectives, we establish an epsilon precision of 0.5 for hydropower revenue, 0.05 for each volumetric reliability to Baltimore, Chester and the Atomic Power Plant, 0.05 for recreational storage reliability and 0.001 for the environmental shortage index as in Giuliani et al. (2014). The Pareto approximate sets, depicted as a 2D projection in panel (c), were obtained across the 25 seeds for each algorithm configuration to have a sense of the performance of the best approximation set for each configuration. Additionally, we obtained individual approximation sets for each random seed to test the reliability of each configuration. The absolute hypervolume was computed for each approximation set (panel d). Each set had its own reference point as described in Section 3.4, which was later normalized to the range [0,1] such that the best possible set was given a hypervolume of 1. Approximation sets with hypervolume near 1 are high-quality, are in close proximity to the best known Pareto front and have captured a diverse representation of the full extent of tradeoffs. Finally, the runtime data was collected every 1000 function evaluations, and hypervolume was computed at each time step to assess runtime performance for every tested Borg MOEA parallel configuration over the 10 h wall-clock time.

Table 1
Default parameters of the Borg MOEA.

Parameters	Value	Parameters	Value
Injection Rate	0.25	PCX Zeta	0.1
SBX Rate	1	UNDX # of Parents	3
SBX Distribution Index	15	UNDX # of Offspring	2
PM Rate	1/L*	UNDX Eta	0.5
PM Distribution Index	20	UNDX Zeta	0.35
DE Crossover Rate	0.1	SPX # of Parents	3
DE Step Size	0.5	SPX # of Offspring	2
UM Rate	1/L*	SPX Epsilon	0.5
PCX # of Parents	3	Number of islands	1,2,4,8,16
PCX # of Offspring	2	Number of cores	32, 64, 128, 256
PCX Eta	0.1	Wall-clock time	10 h

* L = number of decision variables

5. Results

Our diagnostic results explore the balance between exploration, computational demands and the effects of approximation in MC-based evaluations of uncertain objectives. Section 5.1 explores the effectiveness and reliability of each parallel configuration of the Borg MOEA across its tested random seed trials. Effectiveness is explored in terms of the master-worker and multi-master configurations' abilities to attain high quality approximations to the best-known Pareto front (or reference sets). Search reliability refers to the probability that highly effective search is attained. Section 5.2 explores the computational efficiency of the parallel Borg MOEA configurations in solving the LSRB Conowingo reservoir test case in terms of the expected NFE required as

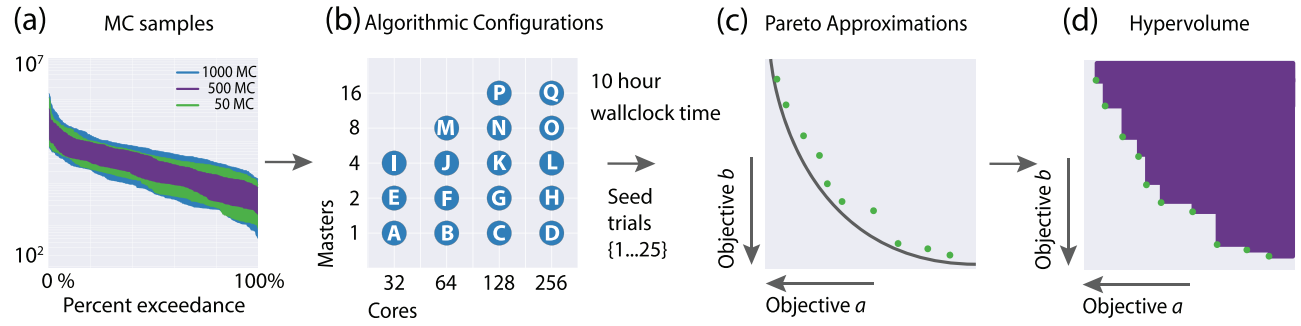

Fig. 7. Schematic of the experimental setup. Each of the three different problem formulations was run under different algorithmic configurations with different core and master counts which were run for 10 h. The Pareto approximation sets were collected across masters for each approximation, and finally the quality of the approximation sets was assessed through hypervolume.

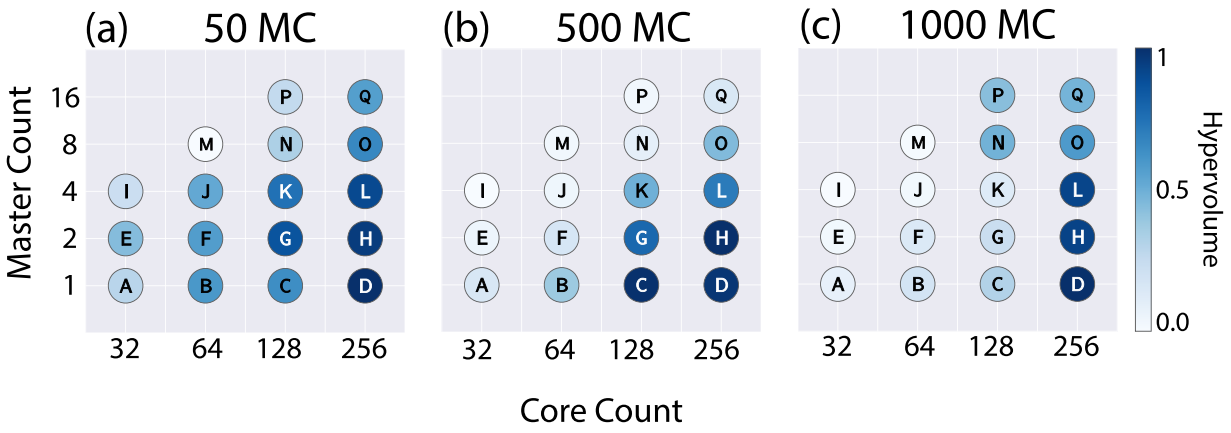

Fig. 8. Best hypervolume (HV) performance for each combination of master and core count across the 25 random seed trials. High HV values are depicted in dark blue and low HV values are depicted light blue. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

well as parallel algorithmic speedups. Section 5.3 explores the decision relevant consequences of the approximate MC evaluations as well as how choosing a parallel configuration of the Borg MOEA can influence tradeoff analyses.

5.1. Effectiveness & reliability of search

Fig. 8 shows the relative hypervolume for the reference sets discovered by each of the algorithmic configurations across their 25 random seed trials. The results were normalized against the best hypervolume attained by each MC sampling level; hence, the parallel Borg MOEA configurations A-Q are only comparable when they exploit the same approximate sampling level (i.e., 50, 500, or 1000 ensemble members). Each circle corresponds to a different parallel configuration of the multi-master Borg, where its location on the x-axis corresponds to its core count and its location on the y-axis to its master count. The color scale represents hypervolume; high performance is depicted in dark blue whereas poor performance is shown in light blue. Panel (a) of Fig. 8 shows the results for the 50 MC sampling scheme. This sampling scheme strongly reduces computational demands at the cost of using highly approximate function evaluations to guide the search. Given that in panel (a) NFE are not as strongly limiting, it is not surprising that 12 out of the 17 configurations achieve at least 50% of the best-known hypervolume for the 50 MC sampling case. Prior studies have consistently shown that providing the Borg MOEA with more NFE strongly enhances its auto-adaptive features and consequently the quality of its search (Hadka and Reed, 2012, 2015).

Transitioning to panel (b) of Fig. 8, the computational demands increase tenfold with the 500 MC sampling scheme. Each function evaluation consists of 500 MC simulations within the same constrained 10 h wall-clock time (i.e., fewer but less approximate function evaluations). In panel (b), the number of successful algorithmic configurations decreases significantly with the reduced NFE available for search. However, high performance is maintained for the 128 and 256 core counts exploiting 1–4 masters.

The importance of the cooperative, auto-adaptive search of the multi-master configurations is illustrated in panel (c) of Fig. 8. The 1000 MC sampling scheme is the most computationally demanding variant of the LSRB test case, as successful algorithmic performance is observed only for the L, H, and D Borg MOEA configurations using 256 cores. It is interesting to note that although configuration D exploits more NFE relative to the L and H configurations, the differences in their relative hypervolume performances are small. This again highlights the potential for the multi-master Borg MOEA to contribute purely algorithmic speedup benefits when compared to the single master-worker configuration using the 256 core counts.

The results in Fig. 8 assume that the best Pareto approximate reference sets for each algorithm configuration are approximated across 25 random search trials. Although often neglected in the water resources literature, the need to use multiple random search trials represents a severe computational cost, especially when exploiting parallel computing systems where users have limited allocations of computing hours. Fig. 9 provides a more detailed analysis of the expected effectiveness and reliability of the alternative parallel configurations of the Borg MOEA if they were run for a single random search trial by displaying the probability of attaining different hypervolume performance thresholds with each multi-master configuration. In this figure each of the columns in panels (a)–(c) correspond to an alternative parallel Borg MOEA configuration. Within each column, the white circle indicates the best overall hypervolume metric value attained by a single trial with that configuration and the color scale indicates the probability that a single trial run with that configuration attains the relative hypervolume value shown on the y-axis. Therefore, a completely dark blue bar filled to the 100% level would indicate ideal performance, designating that a single trial run of an algorithm configuration is both perfectly reliable and effective, always achieving perfect relative hypervolume. The legend in Fig. 9 provides a key for identifying the different parallel Borg MOEA configurations tested. Again, the hypervolumes reported in panels (a)–(c) of Fig. 9 are specific to a given MC sampling scheme.

The attainment results shown in panel (a) of Fig. 9 correspond to the 50 sample MC case (i.e., the fastest but most approximate evaluations). Supporting our prior observations, the fast but approximate function evaluations yield high algorithmic reliability across several parallel Borg MOEA configurations for core counts between 64 and 256. These configurations clearly have performed sufficient NFE during the 10 h wall-clock time for the LSRB test case. It should be noted, however, that there are non-negligible hypervolume differences between each configuration's best single run performances (i.e., white circles) and its reliability attainments (i.e., highest hypervolume with dark blue shading). This highlights that the single best run is often a very misleading result. For example, configuration D (master-worker with a 256 core count) attained the best overall single run for the 50 sample MC results, but the reliability of this master-worker configuration in attaining such high hypervolume performance is extremely low. Moreover, the variance in its performance is high, adding significant uncertainty in how the users would quantify the LSRB test case's tradeoffs. Clearly the attainment plots in Fig. 9 provide a broader context of probabilistic search performance as opposed to the reference set analysis in Fig. 8. If a parallel configuration exhibits outstanding performance in its best single run, but fails to consistently attain high performance, it is of low value to users.

This is especially true for the 500 and 1000 sample MC cases shown in panels (b) and (c) of Fig. 9 respectively; their increased

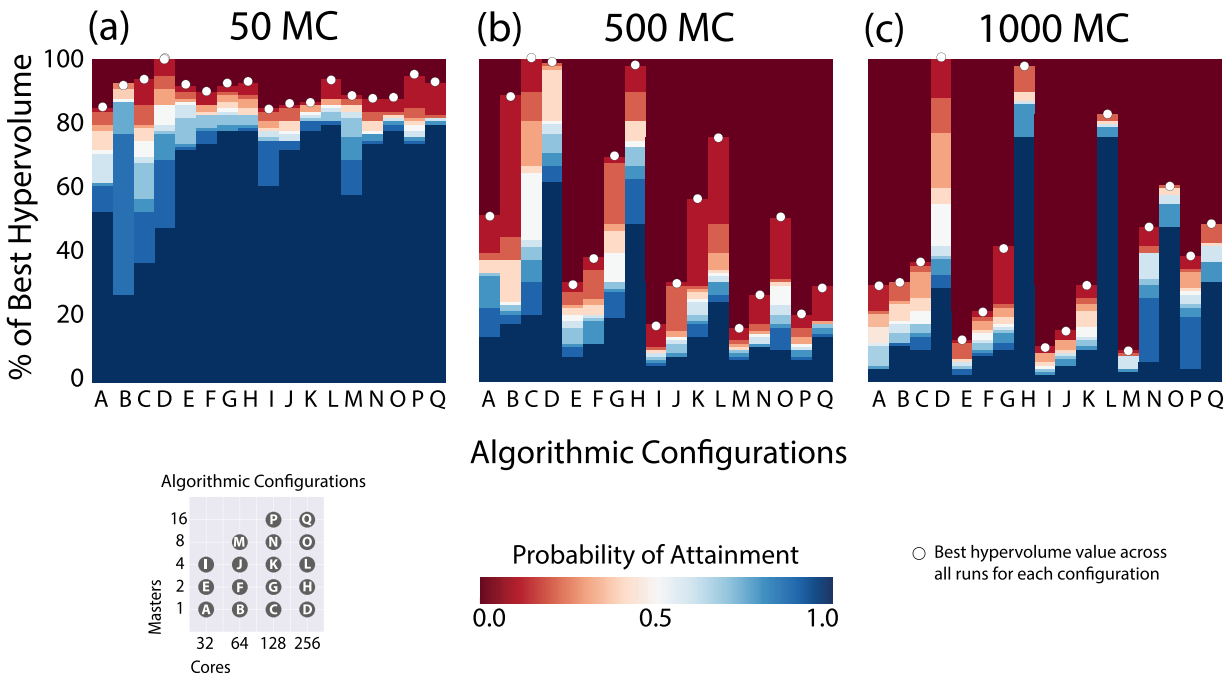

Fig. 9. Best single run and probability of hypervolume attainment. The white dots represent the best overall performance for the algorithmic configuration. The vertical axis shows the percent of the best hypervolume and the color scale shows the probability of attaining a given level of performance. Perfect reliability would be depicted by a blue bar filled to the 100% level with a white dot on top. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

computational demands strongly degrade the performance of most of the tested parallel configurations of the Borg MOEA. Two significant factors shape this result. First, the obvious reduction in the NFE available for search requires dramatically more efficient search pathways for attaining high hypervolume performance. Second, the minimax formulation of the LSRB test case's objectives makes identifying control policies more challenging as more severe droughts are sampled in the 500 and 1000 sample MC cases. In panel (b) of Fig. 9 the best run for the 500 sample MC case is obtained by configurations C (the single master-worker implementation using 128 cores). However, both the master-worker and 2-master variants of the parallel Borg MOEA using 256 compute cores (i.e., configurations D and H in Fig. 9) have high hypervolume attainment probabilities (or best single run reliability). As we further increase the computational demands and problem difficulty by moving to the 1000 sample MC case in panel (c) of Fig. 9, the advantage of the auto-adaptive, interacting search of the multi-master parallelization scheme becomes more evident. Note that for the 256 core count configurations, while the master-worker (configuration D) obtained the best single trial run, the 2 and 4 multi-master configurations (H and L) had a much higher reliability in attaining improved hypervolume performance across all of their individual random trials. As noted in the results of Tang et al. (2007), master-worker parallelization schemes by themselves do not fundamentally change the search dynamics of MOEAs; they simply allow for more NFE per unit wall-clock time, which for relatively easy problems and effective algorithms may be sufficient to attain high levels of performance. However, difficult problems require fundamental changes in the underlying search operations to yield effective and reliable search. There is a broad body of literature on parallel evolutionary search (see review by Alba et al. (2013)) that has shown the island or multi-population parallelization scheme can, when employed well, improve our abilities to solve challenging problems. The cooperative, auto-adaptive search of the 2- and 4-master configurations (H and L) in panel (c) of Fig. 9 is yielding a significant algorithmic benefit when addressing the most challenging instance of the LSRB test case.

5.2. Search efficiency & runtime variability

Building on our attainment results in Fig. 9, this section contributes a more detailed analysis of the runtime search dynamics. Overall, the most reliable search was attained by the 256 core implementations of the single master-worker Borg MOEA as well as the 2- and 4- multi-master configurations (i.e. D, H, and L, respectively). Panels (a)–(c) in Fig. 10 show the runtime hypervolume dynamics for the different MC sampling cases given the same 10 h wall-clock time. Recall that a hypervolume of 1 indicates that a given Pareto approximation has the same hypervolume as the best-known reference Pareto front for its corresponding MC sampling level. The solid lines represent the median hypervolume search progress across the random trials for each of the parallel configurations. The shaded regions bound the 10th and the 90th percentiles of hypervolume performance. As expected, panels (a)–(c) in Fig. 10 show that both the degree of approximation (i.e. the MC sampling level) as well as the parallelization configuration control the NFE that can be performed within the fixed wall-clock time. The 2 and 4-master configurations of the Borg MOEA permit more auto-adaptive cooperation in their search but they allocate fewer cores for performing function evaluations. The cooperative advantage of the multi-master architecture is lost beyond the 4 multi-master count in this case. For instance, for the 8 multi-master configuration with 256 available nodes, the number of worker nodes per master is reduced to 32, as there is an imbalance between communication across masters and the total number of function evaluations for the given wall-clock time. Consequently, the single master-worker will perform more function evaluations relative to the 2 and 4-master instances.

In panel (a) of Fig. 10, the median hypervolume value for the master-worker configuration asymptotically approaches 90% and makes no further improvements throughout the run. The 50 sample MC master-worker configuration performs nearly 5 million function evaluations, but all search beyond 2 million evaluations is a wasted use of computational resources. As was also illustrated in the prior work of Tang et al. (2007), the primary benefit of running the master-worker longer is reducing the variability or uncertainty of search performance across the random trials (i.e., making the results more reliable and less

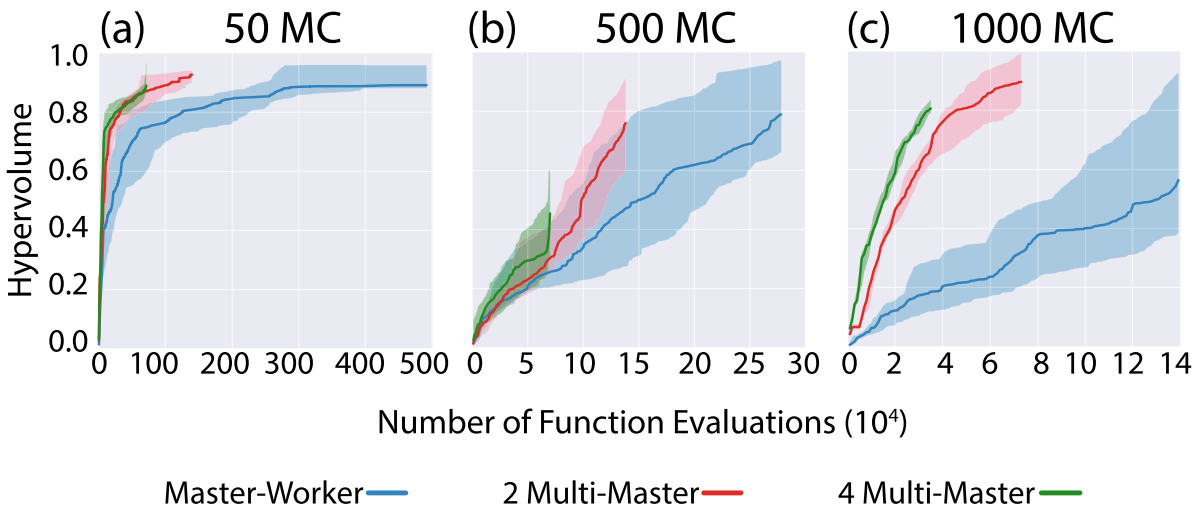

Fig. 10. Runtime hypervolume dynamics for the master-worker, and for the 2 and 4 multi-master configurations. The horizontal axis shows the function evaluations performed by each of the problem formulations during the 10 h wall-clock time; hence the NFE count decreases as the MC sample size increases. The horizontal axis shows the runtime HV normalized against the best HV value at the end of the run for each MC sampling scheme. The solid line represents the mean across random seed trials bounded by the 5*th* and 95*th* percentiles.

sensitive to random initialization). The 2- and 4 -master configurations shown in panel (a) for the 50 sample MC case rapidly achieve higher median hypervolume performance with far fewer NFE, while also significantly reducing random seed variability. These results explain why the 2- and 4-master parallel implementations had more reliable attainment results in Fig. 10 (see configurations H and L in panel (a)).

Panels (b) and (c) in Fig. 10 show a more pronounced improvement in search efficiency and reliability with the multi-master cases. Given the increasing difficulty of transitioning to the 500 and 1000 sampling levels, all of the parallel configurations of the Borg MOEA are forced to exploit far fewer evaluations while seeking to quantify the LSRB test case's tradeoffs. In both the 500 and 1000 sample MC cases, the master-worker configuration has substantially less efficient hypervolume search progress throughout the run, as well as very high random seed variability. For the 500 sample MC case, the upper bound on hypervolume progress of the single master-worker configuration overlaps with the lower bound performance of the multi-master configurations. Put more simply, the best random trials of the master-worker implementation have a similar performance to the worst random trials of the multi-master configurations. Clearly, cooperative, auto-adaptive search across the masters and the controller is providing significant algorithmic search benefits that compensate for their lower NFE. The benefits of coordinated search with masters become even more pronounced for the most challenging 1000 sample MC case. In panel (c) of Fig. 10, the master-worker configuration struggles with the higher difficulty of the 1000 sample MC case, yielding limited hypervolume progress and very high random trial variability. The 2- and 4-master configurations have dramatically lower variability throughout their run and quickly achieve high quality solutions after a substantially lower NFE count. The multi-master Borg MOEA avoids sensitivities to bad initial conditions and stalled search by effectively sharing global knowledge via the controller (see Fig. 5). As has been shown in prior studies, the multi-master Borg MOEA quickly detects and corrects evolutionary search processes without wasting significant computing resources (Hadka and Reed, 2015).

As opposed to typical speedup calculations with a serial algorithm performance baseline, here we use the master-worker configuration of the Borg MOEA as a baseline. For the speedup results shown in Fig. 11 panels (a)–(c), the core count was fixed to 256 and the wall-clock time was constrained to 10 h. Hence, the results from this section reflect the speedup related entirely to differences in algorithmic search dynamics enabled by the multi-master configurations. As discussed in Section 3.4, our speedup calculations account for the NFE needed to attain improving levels of hypervolume performance. For each level of hypervolume, we quantify the ratio between total NFE required by the 2- or 4 -master configurations relative to the NFE required by the master-worker configuration. Any speedups therefore, require that the multi-master configurations attain improved hypervolume performance in less NFE than the master-worker baseline algorithm. Fig. 11 panels (a)–(c) show the worst case hypervolume runtime performance for each algorithm configuration across the different MC sampling cases. In these panels, the dotted lines indicate the zones of hypervolume performance used for the speedup calculations. A broader performance threshold was established for the 50 MC configuration (0–0.9) because the master-worker configuration was able to attain high hypervolume performance. The hypervolume ranges explored for the 500 and 1000 sample MC cases are far more restricted: 0–0.3 and 0–0.35, respectively. A combination of computational constraints on NFE and reduced master-worker performance restricted the comparative ranges where all algorithm configurations had hypervolume performance values that could be compared. Note our analysis focuses on the worst performing seeds because as an algorithm configuration's search reliability decreases, the worst case downside of its uncertain performance is of direct concern.

Analyzing the algorithmic speedups for the 50 and 500 sample MC cases in panel (a) and (b), respectively, of Fig. 11, the 2- and 4-master configurations strongly distinguished themselves at the beginning and end of the runs. Algorithmic speedup at the beginning of the search arises from the increased diversity and early search course corrections of the multi-master configurations. Improved speedups at later stages indicate search progress for the single master-worker implementation of the Borg MOEA reaches diminishing returns far sooner than the multi-master implementations, highlighting the effectiveness of the multi-master coordination in overcoming search stagnation. Transitioning to the 1000 sample MC case shown in panel (c) of Fig. 11, we observe an order of magnitude speedup using the 4-master configuration, i.e. the hypervolume achieved by the 4-master configuration requires 100,000 NFE by the master-worker configuration.

These results show that the multi-master Borg MOEA's parallel architecture is competitive, if not superior, to the master worker configuration, even in the least NFE-constrained 50 sample MC case. As computational and mathematical difficulty increase, its search efficiency and reliability improve dramatically relative to the master worker configuration. Although these insights are in agreement with prior results on very large clusters with 10,000–500,000 compute cores (Hadka et al., 2013; Hadka and Reed, 2015), this study helps to

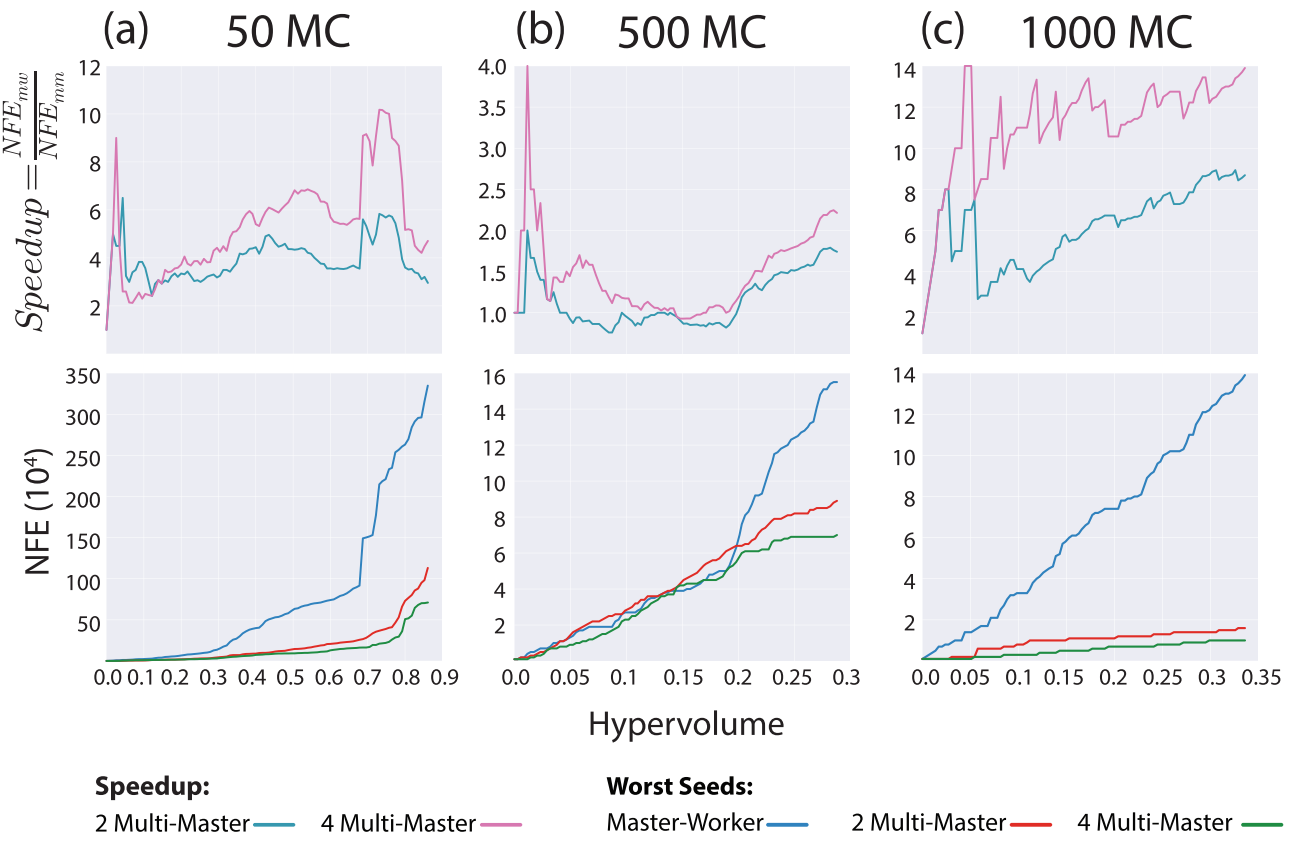

Fig. 11. The algorithmic speedup for the 2 and 4 multi-master configuration is calculated as the ratio of master-worker NFE to multi-master NFE performed within a specified hypervolume threshold. Panels (a)–(c) show the algorithmic speedup for 50 MC–1000 MC, respectively. Depicted below each panel are the worst performing seeds for the mater-worker and the 2 and 4 multi-master used to set the hypervolume threshold for each MC scheme.

generalize our results because the LSRB reservoir control problem is a new and different problem context relative to the problems tested in prior MOEA diagnostic studies, namely in Hadka and Reed (2015) and Reed et al. (2013). Moreover, the relatively small cluster used in this study strongly disadvantages the multi-master Borg MOEA from an NFE perspective, making it even more important for the auto-adaptive, co-operative search features to provide significant algorithmic benefits.

5.3. Consequences of sampling & algorithmic choices

So far, we have compared the performance of algorithmic configurations within each of the MC approximation levels, but not across them. In this section, we explore the decision relevant consequences of users choosing the different levels of approximation in the MC sampling schemes in combination with the alternative parallel Borg MOEA configurations. More simply, we would like to reliably and efficiently capture the best representation of the LSRB test case's operational tradeoffs. In the context of comparing the effects of the different levels of fidelity in estimating the probabilistic evaluations of objectives, it is important too see if the more approximate 50 sample MC scheme yields appropriate representations of the LSRB test case's tradeoffs. For this assessment, we selected the reference sets across all seeds and all the algorithmic configurations for each MC approximation level (i.e., the best known tradeoffs found at each MC sampling level). As discussed in our prior results, the 50 and 500 sample MC configurations have the advantage of being able to perform more NFE compared to the 1000 MC formulation, but at the cost of more approximate function evaluations. These best known LSRB tradeoff solution sets for each MC approximation level were re-evaluated under an independent 1000 sample MC verification set to evaluate the consistency of representations of key tradeoffs. The results of this analysis are shown in Fig. 12.

Panels (a)–(f) of Fig. 12 show the cumulative percent of solutions attaining different performance levels on each of the objectives when evaluated over the optimization and verification MC sets. Objective values over the optimization MC samples are shown with solid lines, and their values over the verification MC samples are shown with dotted lines. The arrows show the direction of preference for each objective. In panel (a) of Fig. 12, which illustrates the relative stability of performance in hydropower revenue, we observe that the 1000 sample MC case yields re-evaluated objective values that are very similar to the optimized values. The verified hydropower performance of the reservoir control policies found using the 50 and 500 MC sample cases, however, show shifts of almost 5 million USD in magnitude for solutions with higher hydropower revenues. Relative to the maximum revenue cases, this deterioration represents a 10% loss in performance in the re-evaluated solutions.

In panel (b) of Fig. 12, the reliability of supplying cooling water to the atomic power plant is stable for the 1000 sample MC case while the 50 and 500 sample MC cases show degradations in performance. The solutions from the 500 sample MC case shows some degradations as high as a 10% reduction in the reliability of supplying cooling water. It should be noted that in the decision context of the LSRB, even a one percent change in reliably providing cooling water to the atomic power plant is concerning; the Susquehanna River Basin Commission has the expressed goal of maintaining cooling water reliability for the atomic power plant near 100%. In panels (c) and (d) of Fig. 12, the water supply objectives for Baltimore and Chester do not shift noticeably with re-evaluation.

In panel (e) of Fig. 12 the reliability of maintaining goal reservoir levels for the recreation objective is limited in the optimized solutions, none of which exceed 72%. Compounding the poor performance in optimization, the solutions from the 50 and 500 sample MC cases degrade even further in re-evaluation. The optimized solutions from the 1000 sample MC case are far more stable in reliably maintaining

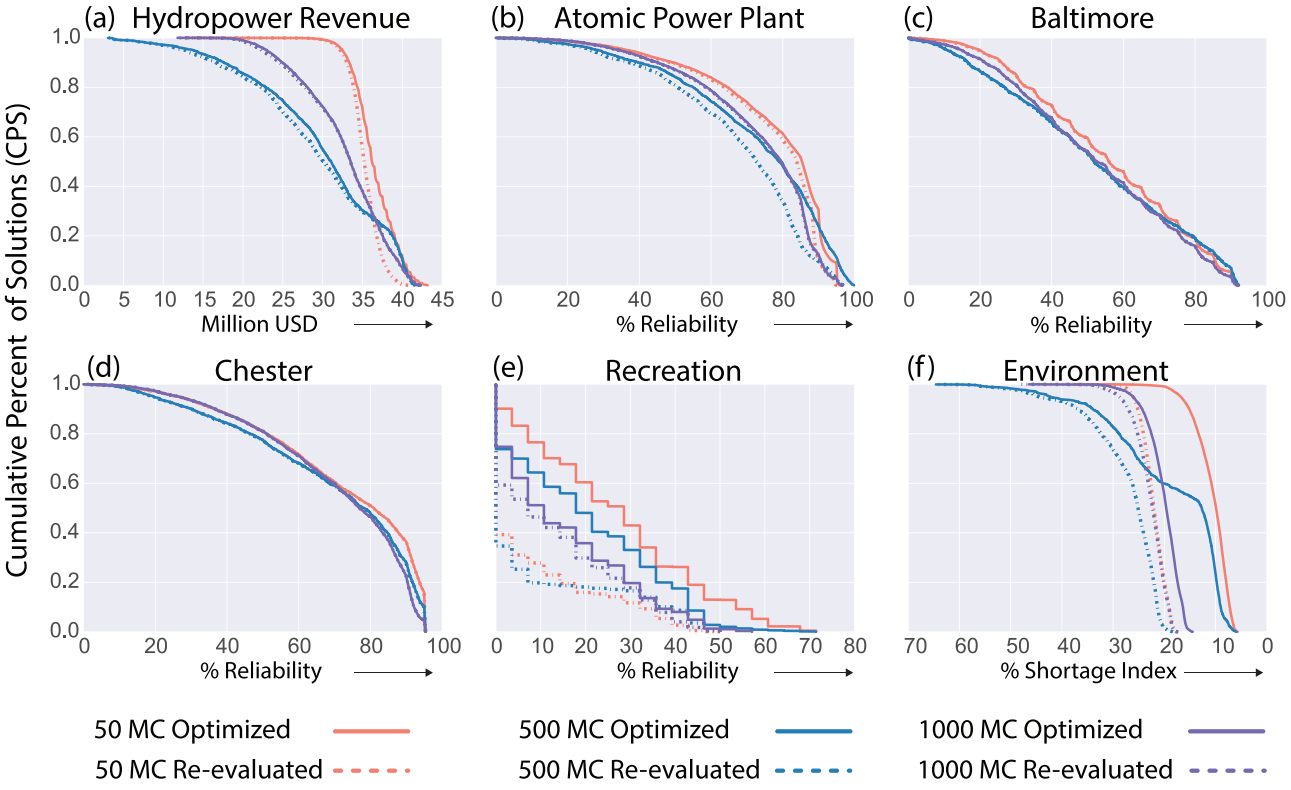

Fig. 12. Cumulative percent of solutions (CPS) for each objective of the LSRB system. All the optimized sets (depicted by solid lines) were re-evaluated under 1000 independent MC samples (dotted lines).

recreation reservoir levels.

Finally, panel (f) of Fig. 12, shows that the federally mandated environmental flows captured in the environmental shortage objective are challenging. None of the best known reference sets achieve a 0% environmental shortage even under the optimization case. The 50 and 500 sample MC cases find solutions that drastically degrade when re-evaluated, with shortages nearly doubling. The solutions found using the more challenging 1000 sample MC evaluations actually highlight the strong potential for shortages (i.e. they have encountered more challenging drought cases). The re-evaluated 1000 sample MC in panel (f) degrades slightly but is overall far more consistent.

Overall, Fig. 12 highlights that for the control policy search context of the LSRB test case, when evaluating the computational tradeoff between more approximate but faster evaluations versus increasing the approximation level of the MC sampling, low sampling rates can yield control policies that do not consistently portray key system tradeoffs. Moreover, these results highlight the importance of verifying that the control tradeoffs presented to stakeholders are stable representations of the system tradeoffs and do not distort decision support negotiations or recommendations (see also Quinn et al. (2017)).

Fig. 13 simulates an interactive decision support exploration of the LSRB test case's key tradeoffs where the stakeholders have expressed two performance requirements: (1) the reliability of supplying cooling water to atomic power plant has to be greater than 95% and (2) deviations from federally mandated environmental flows have to be less than 15%. The major goal in using MOEAs in water resources is to allow this type of interactive *a posteriori* analysis (i.e., brushing or eliminating all solutions that fail to meet these requirements). Often neglected in the literature is that our computational choices (algorithm selection, MC sampling, parallelization, etc.) can distort our perceptions of the tradeoffs themselves. We illustrate this in Fig. 13 by creating a competition across the worst case trial runs of the top performing parallel Borg MOEA configurations such that only those configurations with re-evaluated solutions that meet the cooling water and shortages requirements are allowed to contribute to the plot. We are requiring consistent tradeoff representations that emerge for efficient, effective, and highly reliable search to attain solutions in a zone of high performance with strong objective conflicts. In short, Fig. 13 asks, "What is the worst we can do in understanding our tradeoffs in the context of our requirements?"

In Fig. 13 each of the vertical axes represents an objective where moving upward is always preferred. Hence, the ideal solution would be a straight line that horizontally intersects every axis at its top value. Each of the lines shown in Fig. 13 represents a candidate solution whose re-evaluated performance meets the specified performance requirements. The requirements for the reliability of supplying cooling water to the atomic power plant as well as the environmental flow shortages are shown by brushing these axes to only display solutions meeting the defined acceptable ranges in performance. In Fig. 13, lines intersecting the axes below their top values are showing a reduction or compromise in those objectives. Steep diagonal lines between the axes designate strong tradeoffs. The colors of the lines indicate under which parallel configuration and re-evaluated MC sampling scheme the solutions. The colored pie chart legend shows the percent contribution of the plotted LSRB tradeoff solutions were discovered. Note that only the 2 and 4 master configurations running on 256 compute cores and employing the 1000 realization sampling scheme contributed solutions to Fig. 13. Even for their worst seeds, these configurations were able to reliably find stable solutions for the most challenging case within a limited wall-clock time. Although it may seem somewhat counterintuitive, making the LSRB control problem more mathematically and computationally challenging with the 1000 MC formulation actually improved the multi-master Borg MOEA's value by triggering its auto-adaptive, cooperative search operators.

These results highlight the potential opportunity for future studies to design and evaluate the potential of adaptive strategies for dynamically adjusting the fidelity of design evaluations used to inform search (e.g., the number of MC samples). There are several related

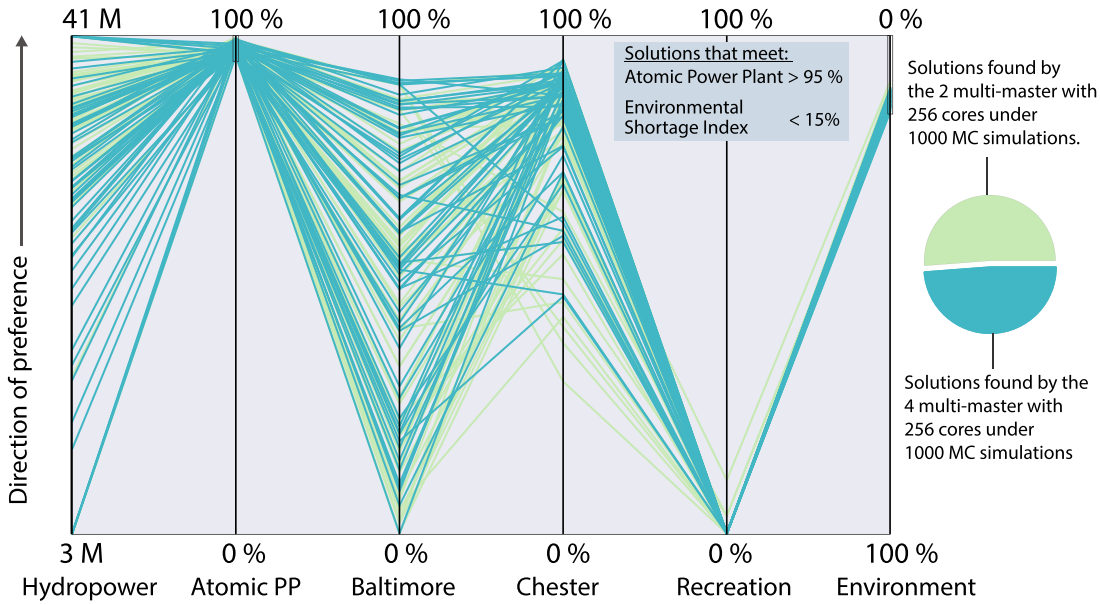

Fig. 13. Parallel axes representation of the LSRB tradeoffs. A horizontal line along the top of the axis would depict ideal performance across all objectives. All the configurations tested in this study were re-evaluated under a common sampling scheme of 1000 MC simulations. The plot shows the solutions that yielded Atomic power plant reliability larger than 95% and Environmental shortage index less than 15%.

bodies of literature that could be valuable for informing extensions of this workincluding dynamic emulation (Castelletti et al., 2012), robust evolutionary optimization (Beyer and Sendhoff, 2007), and cooperative co-evolutionary search (Potter and De Jong, 1994).

6. Conclusions

This study demonstrates the benefits of cooperative parallel MOEA architectures in reliably and effectively finding many objective control policies using Evolutionary Multi-Objective Direct Policy (EMODPS) Search with approximate MC evaluations. Our analysis focuses on the Lower Susquehanna River Basin (LSRB) system where multiple competing objectives for hydropower production, urban water supply, recreation and environmental flows need to be balanced. We consider the tradeoffs between better representing uncertainty in the LSRB system and simplifying this representation to shorten the function evaluation time and allow for greater search when finding Pareto approximate control policies. We used more samples in MC simulations to increase the fidelity and robustness of objective evaluations, at the expense of growing computational demands and increasing the mathematical difficulty of performing well in many more extreme samples. Counter to standard assumptions, simply increasing the number of evaluations does not necessarily improve the reliability of finding high quality and stable representations of key tradeoffs. This study demonstrated that increasing MC sampling in combination with auto-adaptive, coordinated parallel search can dramatically enhance the efficiency and reliability of search while better representing SRB tradeoffs.

Overall, we observed that the 2 and 4 island versions of the multi-master Borg MOEA solved the problem with the highest-quality results within a 10-hour wall-clock time with 256 cores. blueWhile the master-worker configuration can be effective when maximizing the core count and when solving problems that are known to be easy, when solving a problem with higher complexity (e.g. with multiple MC simulations, many objectives and high computational demands) or when seeking to minimize the need for replicate random seeds trials, we would recommend using multi-master Borg MOEA implementations because of their increased reliability for difficult problems. Our measure of algorithmic speedup demonstrated that these multi-master configurations show high benefits entirely due to cooperative search between islands. This avoids the need for running large numbers of replicate random trials; hence, we can employ less computational effort to account for MOEAs' stochastic use of operators and random initial populations. The computational efforts can instead be employed toward improving model accuracy and better representing hydro-climatic uncertainties or evaluating other sources of uncertainties. These algorithmic configurations also proved to have crucial impacts in the decision space, as they discovered high quality tradeoffs for critical objectives, such as the Atomic Power Plant reliability, for which a one percent reduction in performance can seriously threaten the safety of the system.

This study is in agreement with a long history of work (e.g., the recent review by Alba et al. (2013), also in Hadka and Reed (2015)) that shows hierarchical hybrid parallelization schemes like the multi-master schemes shown here fundamentally increase the mathematical difficulty of applications that can be addressed efficiently and reliably. This work shows that the multi-master implementations of the Borg MOEA generalize its auto-adaptivity and add cooperative search operators that allow users to reliably solve complex water systems problems. The algorithmic benefits of increased reliability on harder problems make it feasible to better incorporate uncertainties within the search. To generalize our understanding of the potential benefits of the multi-master schemes' algorithmic speedups and their abilities to reliably provide high quality representations of key reservoir tradeoffs, it would be beneficial to extend the analyses in this study to a diverse set of river basins and other multi-sectoral decisions contexts.

The key insights of this study are useful for water managers determining how to optimize operating policies in river basin systems with multiple reservoirs and uncertain inputs impacting their operations In these systems, EMODPS solution strategies require MOEAs that are highly scalable and extensible to emerging parallel computing architectures, and our results indicate that algorithms with multi-population architectures best meet these needs. These algorithms will become even more useful as the availability of multiple cores in desktop, laptop, and even smartphones platforms continue to grow and increase opportunities to utilize parallel metaheuristics in water resources applications.

Acknowledgements

Portions of this work were supported by the National Science Foundation through the Network for Sustainable Climate Risk

Management (SCRiM) under NSF cooperative agreement GEO-1240507 as well as the Consejo Nacional de Ciencia y Tecnologia (CONACYT) Fellowship no. 313591. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the US National Science Foundation or CONACYT.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.advwatres.2017.09.014.

References

Alba, E., Luque, G., Nesmachnow, S., 2013. Parallel metaheuristics: recent advances and new trends. Int. Trans. Oper. Res. 20 (1), 1–48.
Beyer, H.-G., Sendhoff, B., 2007. Robust optimization–a comprehensive survey. Comput. Methods Appl. Mech. Eng. 196 (33), 3190–3218.
Cantu-Paz, E., 2000. Efficient and accurate parallel genetic algorithms. 1 Springer Science & Business Media.
Castelletti, A., Galelli, S., Ratto, M., Soncini-Sessa, R., Young, P.C., 2012. A general framework for dynamic emulation modelling in environmental problems. Environ. Modell. Softw. 34, 5–18.
Coello, C.A.C., Lamont, G.B., Van Veldhuizen, D.A., et al., 2007. Evolutionary algorithms for solving multi-objective problems. 5 Springer.
Cohon, J.L., Marks, D.H., 1975. A review and evaluation of multiobjective programing techniques. Water Resour. Res. 11 (2), 208–220.
Crowl, L.A., 1994. How to measure, present, and compare parallel performance. IEEE Parallel Distrib. Technol. 2 (1), 9–25.
Deb, K., Gupta, H., 2006. Introducing robustness in multi-objective optimization. Evol. Comput. 14 (4), 463–494.
Giuliani, M., Castelletti, A., Pianosi, F., Mason, E., Reed, P.M., 2015. Curses, tradeoffs, and scalable management: advancing evolutionary multiobjective direct policy search to improve water reservoir operations. J. Water Resour. Plann. Manage. 142 (2), 04015050.
Giuliani, M., Herman, J.D., Castelletti, A., Reed, P.M., 2014. Many-objective reservoir policy identification and refinement to reduce policy inertia and myopia in water management. Water Resour. Res. 50 (4), 3355–3377.
Giuliani, M., Pianosi, F., Castelletti, A., 2015. Making the most of data: an information selection and assessment framework to improve water systems operations. Water Resour. Res. 51 (11), 9073–9093.
Giuliani, M., Quinn, J.D., Herman, J.D., Castelletti, A., Reed, P.M., 2017. Scalable multiobjective control for large scale water resources systems under uncertainty. IEEE Trans. Control Syst. Technol. doi:http://dx.doi.org/10.1109/TCST.2017.2705162.
Goldberg, D., 2002. The Design of Innovation: Lessons from and for Competent Genetic Algorithms. kluwer academic publishers, Norwell, Mass.
Gopalakrishnan, G., Minsker, B.S., Goldberg, D.E., 2003. Optimal sampling in a noisy genetic algorithm for risk-based remediation design. J. Hydroinf. 5 (1), 11–25.
Guariso, G., Rinaldi, S., Soncini-Sessa, R., 1986. The management of lake como: a multiobjective analysis. Water Resour. Res. 22 (2), 109–120.
Hadka, D., Madduri, K., Reed, P., 2013. Scalability analysis of the asynchronous, master-slave borg multiobjective evolutionary algorithm. Parallel and Distributed Processing Symposium Workshops & PhD Forum (IPDPSW), 2013 IEEE 27th International. IEEE, pp. 425–434.
Hadka, D., Reed, P., 2012. Diagnostic assessment of search controls and failure modes in many-objective evolutionary optimization. Evol. Comput. 20 (3), 423–452.
Hadka, D., Reed, P., 2013. Borg: an auto-adaptive many-objective evolutionary computing framework. Evol. Comput. 21 (2), 231–259.
Hadka, D., Reed, P., 2015. Large-scale parallelization of the borg multiobjective evolutionary algorithm to enhance the management of complex environmental systems. Environ. Modell. Softw. 69, 353–369.
Haimes, Y.Y., Hall, W.A., 1977. Sensitivity, responsivity, stability and irreversibility as multiple objectives in civil systems. Adv. Water Resour. 1 (2), 71–81.
Hamarat, C., Kwakkel, J.H., Pruyt, E., Loonen, E.T., 2014. An exploratory approach for adaptive policymaking by using multi-objective robust optimization. Simul. Modell. Pract. Theory 46, 25–39.
Hashimoto, T., Loucks, D.P., Stedinger, J.R., 1982. Reliability, resiliency, robustness, and vulnerability criteria for water resource systems. Water Resour. Res. 18 (1).
Herman, J.D., Zeff, H.B., Lamontagne, J.R., Reed, P.M., Characklis, G.W., 2016. Synthetic drought scenario generation to support bottom-up water supply vulnerability assessments. J. Water Resour. Plann. Manage. 142 (11), 04016050.
Kasprzyk, J.R., Nataraj, S., Reed, P.M., Lempert, R.J., 2013. Many objective robust decision making for complex environmental systems undergoing change. Environ. Modell. Softw. 42, 55–71.
Kirsch, B.R., Characklis, G.W., Zeff, H.B., 2012. Evaluating the impact of alternative hydro-climate scenarios on transfer agreements: practical improvement for generating synthetic streamflows. J. Water Resour. Plann. Manage. 139 (4), 396–406.
Knowles, J., Corne, D., 2002. On metrics for comparing nondominated sets. Evolutionary Computation, 2002. CEC'02. Proceedings of the 2002 Congress on. 1. IEEE, pp. 711–716.
Koutsoyiannis, D., Economou, A., 2003. Evaluation of the parameterization-simulation-optimization approach for the control of reservoir systems. Water Resour. Res. 39 (6).
Laumanns, M., Thiele, L., Deb, K., Zitzler, E., 2002. Combining convergence and diversity in evolutionary multiobjective optimization. Evol. Comput. 10 (3), 263–282.
Loucks, D.P., Van Beek, E., Stedinger, J.R., Dijkman, J.P., Villars, M.T., 2005. Water Resources Systems Planning and Management: An Introduction to Methods, Models and Applications. Paris: Unesco.
Maier, H.R., Kapelan, Z., Kasprzyk, J., Kollat, J., Matott, L., Cunha, M., Dandy, G.C., Gibbs, M.S., Keedwell, E., Marchi, A., et al., 2014. Evolutionary algorithms and other metaheuristics in water resources: current status, research challenges and future directions. Environ. Modell. Softw. 62, 271–299.
Mortazavi-Naeini, M., Kuczera, G., Kiem, A.S., Cui, L., Henley, B., Berghout, B., Turner, E., 2015. Robust optimization to secure urban bulk water supply against extreme drought and uncertain climate change. Environ. Modell. Softw. 69, 437–451.
Moss, R., Fisher-Vanden, K., Delgado, A., Backhaus, S., Barrett, C., Bhaduri, B., Kraucunas, I., Reed, P.M., Rice, J., Sue Wing, I., Tebaldi, C., 2016. Understanding dynamics and resilience in complex interdependent systems. Global Change Research Program Interagency Group on Integrative Modeling: 38, Washington, D.C., U.S.
Müller, R., Schütze, N., 2016. Multi-objective optimization of multi-purpose multi-reservoir systems under high reliability constraints. Environ. Earth Sci. 75 (18), 1278.
Nicklow, J., Reed, P., Savic, D., Dessalegne, T., Harrell, L., Chan-Hilton, A., Karamouz, M., Minsker, B., Ostfeld, A., Singh, A., et al., 2009. State of the art for genetic algorithms and beyond in water resources planning and management. J. Water Resour. Plann. Manage. 136 (4), 412–432.
Nowak, K., Prairie, J., Rajagopalan, B., Lall, U., 2010. A nonparametric stochastic approach for multisite disaggregation of annual to daily streamflow. Water Resour. Res. 46 (8).
Oliveira, R., Loucks, D.P., 1997. Operating rules for multireservoir systems. Water Resour. Res. 33 (4), 839–852.
Potter, M.A., De Jong, K.A., 1994. A cooperative co-evolutionary approach to function optimization. International Conference on Parallel Problem Solving from Nature. Springer, pp. 249–257.
Quinn, J.D., Reed, P.M., Giuliani, M., Castelletti, A., 2017. Rival framings: a framework for discovering how problem formulation uncertainties shape risk management tradeoffs in water resources systems. (In Review) Water Resour. Res.
Randall, D., Cleland, L., Kuehne, C.S., Link, G.W., Sheer, D.P., 1997. Water supply planning simulation model using mixed-integer linear programming "engine". J. Water Resour. Plann. Manage. 123 (2), 116–124.
Reed, P.M., Hadka, D., 2014. Evolving many-objective water management to exploit exascale computing. Water Resour. Res. 50 (10), 8367–8373.
Reed, P.M., Hadka, D., Herman, J.D., Kasprzyk, J.R., Kollat, J.B., 2013. Evolutionary multiobjective optimization in water resources: the past, present, and future. Adv. Water Resour. 51, 438–456.
Rosenstein, M.T., Barto, A.G., 2001. Robot weightlifting by direct policy search. International Joint Conference on Artificial Intelligence. 17. Citeseer, pp. 839–846.
Sheer, D.P., Dehoff, A., 2009. Science-based collaboration: finding better ways to operate the conowingo pond. Am. Water Works Assoc. J. 101 (6), 20.
Singh, A., Minsker, B.S., 2008. Uncertainty-based multiobjective optimization of groundwater remediation design. Water Resour. Res. 44 (2).
Soncini-Sessa, R., Weber, E., Castelletti, A., 2007. Integrated and Participatory Water Resources Management-Theory. 1 Elsevier.
Swartz, P., 2006. Conowingo Pond Management Plan. Technical Report 242. Susquehanna River Basin Commission.
Tang, Y., Reed, P.M., Kollat, J.B., 2007. Parallelization strategies for rapid and robust evolutionary multiobjective optimization in water resources applications. Adv. Water Resour. 30 (3), 335–353.
Tsoukalas, I., Kossieris, P., Efstratiadis, A., Makropoulos, C., 2016. Surrogate-enhanced evolutionary annealing simplex algorithm for effective and efficient optimization of water resources problems on a budget. Environ. Modell. Softw. 77, 122–142.
Ward, V., 2015. Confronting tipping points: How well can multi-objective evolutionary algorithms support the management of environmental thresholds.
Zatarain Salazar, J., Reed, P.M., Herman, J.D., Giuliani, M., Castelletti, A., 2016. A diagnostic assessment of evolutionary algorithms for multi-objective surface water reservoir control. Adv. Water Resour. 92, 172–185.
Zhang, J., Wang, X., Liu, P., Lei, X., Li, Z., Gong, W., Duan, Q., Wang, H., 2017. Assessing the weighted multi-objective adaptive surrogate model optimization to derive large-scale reservoir operating rules with sensitivity analysis. J. Hydrol. 544, 613–627.
Zitzler, E., Thiele, L., Laumanns, M., Fonseca, C.M., Da Fonseca, V.G., 2003. Performance assessment of multiobjective optimizers: an analysis and review. IEEE Trans. Evol. Comput. 7 (2), 117–132.